Boehringer Ingelheim

Boehringer Ingelheim
Pharmaceuticals Inc.

Dockets Management Branch (HFA-305)
Food and Drug Administration
5630 Fishers Lane, Room 1061
Rockville, MD 20852

December 10, 2004

Docket No. 2004D–0377, International Conference on Harmonisation; Draft Guidance on E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs

Heidi C. Reidies
Director
Drug Regulatory Affairs
Telephone (203) 798-4345
Telefax (203) 791-6262
E-Mail hreidies@rdg.boehringer-ingelheim.com

900 Ridgebury Rd/P.O. Box 368
Ridgefield, CT 06877-0368
Telephone (203) 798-9988

Dear Sir or Madam:

Boehringer Ingelheim appreciates the opportunity to give comments on the subject draft guidance. Our comments are provided on the following pages, identified by section number/title of the guidance and line number where appropriate.

Please contact the undersigned with any questions or comments on this correspondence.

Sincerely,

Heidi C. Reidies
Drug Regulatory Affairs

2004D-0377 C4

Boehringer Ingelheim comments on:

ICH E14 Guideline "The Clinical evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for non-Antiarrhythmic Drugs - Step 2"

General Comments

1. **Comment concerning the extent to which negative non-clinical studies (see ICH SB7 guidance) can exclude a clinical risk beyond reasonable doubt**

   The currently used preclinical testing strategy is considered to be highly sensitive for detecting drugs with a proarrhythmic potential, at least by those who conduct such studies. Admittedly, each test has certain pitfalls, but together they have been proven to detect all known compounds with QT-prolonging and proarrhythmic potential. The ICH S7b guideline concentrates on the establishment of an integrated risk assessment such that all relevant data are used to define risk. This is meant to compensate for the fact that a given test may not detect all drugs that can affect ventricular repolarization; however, all known repolarization-inhibitors will prove positive in at least one of the commonly employed test systems (hERG, AP or in vivo ECG).

   In the case where all pre-clinical studies (according to ICH S7b) are negative, the clinical risk for QT prolongation should be considered to be extremely low, particularly in the absence of an active human metabolite. For compounds with a clear negative pre-clinical signal, if no other relevant QT prolongation (individually and in the population studied) has been observed during the early clinical trials, a 'thorough QT/QTc study' could be postponed until closer in time to the final regulatory submission without any appreciable increase in clinical risk. (Such an approach would lead to standard ECG monitoring in phase III appropriate to the therapeutic area for the drug.) In those cases where the 'thorough QT/QTc study' is deferred to a later time, a meta-analysis of ECG data from completed clinical trials could help provide guidance on whether a 'thorough QT/QTc study' is necessary or not. In some situations, it might be possible to forego the need for a 'thorough QT/QTc study' if a meta-analysis of the observed data of all clinical trials of the investigational drug (in which the protocol has defined the recording and quantitative measurement of ECGs) provides negative results. Such a meta-analysis should be based on well-defined conditions that merit poolability across trials (e.g., ECG measurements that cover the dosing interval or at least C_{max}), address aspects that would normally contribute to the design of a 'thorough QT/QTc study' (e.g., inclusion of therapeutic and supratherapeutic doses, adequate sensitivity as provided by an active control), and take into account the potential heterogeneity across trials (see Webber (2003)).

   If not all pre-clinical studies are negative or if any relevant QT prolongation is observed in early clinical trials despite negative results of the pre-clinical studies, a 'thorough QT/QTc study' should be performed with an investigational agent as soon as its characteristics, metabolites and nature of interactions with other compounds is known.

   The goal with such an approach (to postpone the 'thorough QT study' in the absence of preclinical and clinical QT signals until Phase III) would be to minimize the clinical risk to an acceptable level.

2. **Comment concerning the categories of drugs for which there would be no need for a clinical thorough QT/QTc study.**

Usually approved drugs with comparable rates between treated and control patients of certain adverse events (like torsades or sudden death) with no other QT liability should not require a 'thorough QT/QTc study' (as QT is only a surrogate for torsades). If a new indication is sought for such a drug in the same population with the same or lower doses, there should be no need to perform a 'thorough QT/QTc study'. If a new indication is sought in a different population with the same or lower doses and there are no drug-drug interactions of concern from a QT prolongation perspective, there should be no need to perform a 'thorough QT/QTc study'.

Categories of drugs for which there should be no need for a 'thorough QT/QTc study' should include drugs for life-threatening illnesses (e.g. in oncology or virology etc). Other drugs for treatment with certain clinical need or drugs for single administration (in contrast to chronic treatment) might be evaluated using criteria different from those described in the draft guidance (e.g., higher acceptance range for the non-inferiority margin or higher thresholds for risk categorizations). This might include a 'thorough QT/QTc study' in patients instead of subjects. In any of these cases, a discussion with the authorities would be recommended.

3. **Comment concerning footnote # 2 on the definition of a negative 'thorough QT/QTc study' as on where the largest time-matched mean difference between the drug and placebo (baseline-subtracted) for QTc intervals around 5ms or less, with a one-sided 95% confidence interval that excludes an effect >8.0ms. This upper bound was chosen to reflect the uncertainty related to the variability of repeated measurements.**

Based on the observed variability, 10 ms seems more appropriate as the upper limit of the confidence interval.

The "largest time-matched mean difference between drug and placebo (baseline subtracted)" is an ambiguous specification of the endpoint and the analytical method. There are several mathematical interpretations of this description which lead to different statistical properties of the endpoints and their analyses. One could interpret this endpoint as a within-subject calculation (since endpoints are subject-based) which includes the random variability (e.g., an unusual placebo value could drive selection of the largest within-subject difference).

A preferred interpretation would be "the maximum over time of the one-sided 95% confidence intervals of the time-matched (mean) differences between drug and placebo (baseline subtracted)", based on a pre-specified time interval with strong relation to the pharmacokinetic time course of the investigational drug. This method preserves the alpha level of the global test since it is a closed testing procedure and has the advantage that it does not mix the observations of the various time points (which can produce biased estimates of effect).

An important question which should be considered is the number of defined sampling times for ECG recordings, which might depend on the pharmacokinetic profile of the investigational drug. The larger the number of time points, the larger will be the expectation and variability of the maximum time-matched difference and, consequently,

the smaller will be the power. Reduced power is not an issue if focus is on the mean rather than the maximum. A potential weakness of the mean is that it could have the effect of flattening out any QT/QTc effect that might otherwise be apparent with the maximum. If the time interval is, however, prespecified and related to peak plasma levels, then the maximum and the (clinically relevant) mean are closely related to each other.

The definition of the non-inferiority margin should be defined strictly on the basis of known clinical risk (cf. criteria of item 3 above). An upper margin of 10 ms of the one-sided 95% confidence interval has been deemed reasonable by regulatory agencies for several completed and ongoing 'thorough QT/QTc studies'.

Alternatively, Holter ECG recordings could be considered which might require different analysis methods (and non-inferiority margins or limits), but which might help identify potential reasons of QT individual prolongations.

4. **Comment concerning footnote #3 on the categorization of clinical risk for drugs that prolong the mean QT/QTc interval by around 5ms or less, 6 to 10ms, 11 to 15ms, 16 to 20ms and those that prolong the mean QT/QTc interval by more than 21ms.**

The minimum threshold of 5ms should be raised to 10ms.

A categorization of the clinical risk based on observed mean prolongation by larger than certain margins would allow one to weigh the risks and benefits of the drug within its therapeutic area. The dilemma is that these category limits imply a more linear relationship between QT interval prolongation and clinical risk for torsade de pointes and similar events than is borne out by historical data at the current time, and this makes them unsatisfactory for anything other than descriptive purposes. For example, Prof. Malik has commented at DIA meetings that placebo can have a mean increase of 10 ms. If the limits for these categories are intended to relate to the degree of clinical risk, they should be set on the basis of data for drugs with potential QT prolongation and risk for torsade de pointes that use the methodology currently proposed for 'thorough QT/QTc studies" (e.g., central ECG reading, ECG frequency and timing).

Another important aspect is that the interpretation of the different limits might be different for different therapeutic areas (e.g., life-threatening illness vs. hypertension) or based on the medical need of the investigational drug.

If this categorization is retained in the guidance, it would be helpful to clarify in the guidance that this categorization reflects placebo-corrected differences.

5. **Comment concerning section 3.0 on the relative emphasis on population mean value versus individual outlier analysis in determining the outcome of the "thorough QT/QTc study" as either positive or negative.**

Outlier analysis as a primary analysis may completely change sample size. This also means that the quality of each individual ECG becomes crucial and may lead to the need for an even larger number of replicate ECGs, individual heart rate corrections, etc.

As the variance of the maximum of a set of identically distributed random variables is typically larger than the variance of these variables itself, the primary emphasis should be the population mean value of the endpoints, and the individual outlier analysis should be considered secondary.

For the outlier analysis of QTc, the effect of the applied heart rate correction formula is crucial, as it can be the sole influential factor in making a QTc value or change into an outlier.

6. **Comment on the extent to which results of a negative "clinical thorough QT/QTc study" can be extrapolated to exclude a risk in patients, especially in the context of patients with increased risk (e.g. extending the indication of an antihypertensive drug to include subsequently those with chronic heart failure).**

   To our knowledge, there has been no example where a QT signal of a torsades-causing drug could not be detected in healthy volunteers. This is an argument to stay with this extrapolation.

   From another perspective, extrapolation should be based on the legitimacy of that extrapolation. But, how does one gather evidence to support the extrapolation, and what kind of extrapolation would make sense? The pharmacologic properties of the drug (including the nature of drug-drug interactions) and doses to be used in the more fragile patient population would seem to dictate whether extrapolation of no anticipated increased risk might make sense (cf. point 2).
   \

7. **Comment concerning the sensitivity requirements for active controls**
   In the current draft guideline it is not clearly advised as to what extent an active control must show assay sensitivity and what timing for the ECG measurements should be used. It should be mentioned that a single administration of the active control is deemed to be sufficient. Logistically, one should avoid having the active control (such as moxifloxacin) be an integral treatment in the 'thorough QT/QTc study" design because it is not efficient.

   For the timing of ECG recordings for the active control it seems to be important that it matches the pharmacokinetic profile of this active control and <u>not</u> necessarily mirror the timing for the ECG recordings for the investigational drug.

8. **Comment concerning methods of QT correction**
   Current methods of the QT-correction for standard 10-sec ECG recordings do not account for natural heart rate variability. To overcome the most pronounced effects of heart rate variability it should be recommended that the QT correction be performed not on the QT/RR relation of each wave form, but rather on the QT/RR relation within the 10 sec recording.

Also, different people might have different heart rate variability. It would be helpful to know if heart rate limits can be defined that identify ECG measurements as being unaffected by hysteresis.

Specific Comments:

1. **Section 2.1.1**
   Exclusion criteria and criteria for the discontinuation of individual subjects: This point obviously aims at the safety of patients/volunteers participating in clinical trials. If it becomes effective one might be obliged in a number of projects to conduct the thorough QT study very early as otherwise one would not be able to include the target patient population (heart failure etc). The suggested rules for the discontinuation of individual patients /subject, mainly the increase in QT(c) interval would require a careful "online analysis" of the ECG/QT data. Manual measurement could be used rapidly; it is however not very precise and time consuming. It is also not common practice to record the measured QT / QTcB / QTcF interval in a spreadsheet (required for the detection of changes). For logistical reasons there is usually a lag time until the QT data generated by the core labs are available. Our suggestion is to skip this as a general requirement for clinical studies as it is not practically applicable or at least limit it to certain exceptions of increased risk.

2. **Section 2.1.2**
   It should be avoided that the guidance imposes a thorough QT-study too early, when (1) we still have not enough data to design a really "definitive" trial, and (2) at a stage where there is still considerable uncertainty about the continuation of the development due to other reasons than QT.
   Alternate suggestion to *Line 229 to 232: "The thorough QT-study would typically be conducted early in clinical development... "*
   "The timing of a thorough QT-study in clinical development would typically depend on a risk-assessment based on preclinical data and data from early clinical pharmacology studies. If no or low risk is indicated from preclinical data, a thorough QT-study can be performed in late clinical development, if such a trial is necessary at all. A preclinical QT-liability, however, might suggest an earlier QT-study. Also evaluation of ECGs recorded during other well-controlled clinical pharmacology trials (e.g. rising dose studies) may allow sufficient risk assessment and postpone the need for a thorough QT-study into a later stage of development, especially if no sign of QT-prolongation are seen from these ECGs. In any case, basic clinical data, including tolerability and pharmacokinetics, are necessary for the design and conduct of a thorough QT-study."

3. **Section 2.1.3**
   The draft guidance (section 2.1.3) states "*If the 'thorough QT/QTc study' is positive, additional evaluation in subsequent clinical studies should be performed.*" One objective of this evaluation should be to fully characterize the dose-, concentration-, and time-relationships of the drug on the QT/QTc interval in the target patient population(s) at therapeutic and supratherapeutic serum concentrations. The latter can be achieved in two ways: through administration of high doses or use of metabolic inhibitors (if applicable).

4. **Section 2.2.2 - Line 249 etc**
   The sentence *"As this study has a critical role..."* should be deleted. In most cases (i.e. no preclinical liability) the thorough QT-study is mainly for approval, not for guiding the clinical development.

5. **Section 2.2.2**
   The statements for the accuracy and precision of the QT measurements are not sufficient for the evaluation of the reproducibility of the measurements with other methods. It is well known that the manual measurements of different readers might differ from each other. This should be addressed in a specific section in the guidance.

6. **Section 2.2.2 – Line 348**
   The draft guidance states that "*Readers of ECGs should be blinded to time, treatment and subject identifier, and one reader should read all the ECG recordings from a given subject.*" However, estimates of variability may have been based on the practice of having central ECG readers evaluate ECGs within a patient serially without knowledge of time or treatment. Dr. Morganroth has mentioned this previously at DIA ECG meetings, and truly blinded reading of ECGs would increase the variability substantially. The choice of non-inferiority margin (e.g., 8 or10 ms) might be historically based on this partially blinded reading methodology that is not specified in the guidance. So, the guidance should reflect caution regarding the specification of statistical bounds or estimates of variability commensurate with the quality of the methodology upon which such limits are based.

   Also, if the intent is to identify worsening of morphological abnormalities, one needs to review ECGs relative to baseline conditions. This would have to be done with a separate process that is not described in the guidance (i.e., central ECG reader could not be blinded to patient or time), potentially in addition to the blinded reading.

7. **Section 2.2.2 – Line 350**
   The draft guidance suggests having a subset of normal and abnormal data reread in a blinded manner to assess inter- and intra-reader variability. However, the value of such an assessment is unclear. On the other hand, such methodology where re-assessments are performed could serve to validate semi-automated or even fully-automated technologies.

8. **Section 3.1**
   The draft guidance puts emphasis on the evaluation of time-matched ECG recordings. While this methodology is recognised to overcome artefacts due to potential effects of the circadian rhythm, it also adds additional variability to the measurements. Hence, at the trial design stage, the impact of both effects should be weighed and a proposal for an alternative method for baseline correction is optional.

9. **Section 3.1 – Line 406**

Comment on line 406, section 3.1: Early trials are not designed to detect relatively small effects in QT/QTc prolongation, therefore we suggest removing the descriptor "early". Similarly, on line 409, we suggest removing the descriptor "late".

10. **Section 3.1.2 – Line 450**
    The draft guidance (section 3.1.2 – line 450) states that "*care should be taken to exclude ECG recordings collected during times of rapid heart rate changes due to this QT/RR hysteresis effect*". With ECG measurements captured within 10 sec recordings traditionally, there appears to be no way to exclude measurements affected by hysteresis. Should the standard length of a recording be extended to, say, 30 ms? Should a stable heart rate be established before ECG measurements are taken? In such a case, an online measurement for heart rate would be needed, or Holter ECG recording should be considered.

11. **Section 3.1.2**
    In the current version of the text, the last sentence of Section 3.1.2 should be moved to Section 2.2.2, as it explains the recording of the data and not the applied correction formula.

12. **Section 3.1.2**
    In Section 3.1.2, it might be recommended to report the following parameters as at least secondary parameters in all trials where ECGs have been recorded and measured: the QT interval and the heart rate corrected parameters QTcF and QTcB. These parameters would allow an indirect (maybe historical) comparison of the magnitude of an effect, even if the study population is not sufficiently characterized by these correction formulas.

    In item 3 of this section, it is preferred to determine study specific QT-RR relations and the resulting QT corrections for all linear and nonlinear regression techniques only on drug-free baseline data, not on the on-treatment measurements of placebo treatment.

    In item 4 of this section, the parabolic regression formulas $QTc=QT/RR^a$ with *estimated* coefficient a should be explicitly specified, as this is a common technique to assess the QT-RR relation of the study population (commonly QTcN) or the individuals (QTcI).

13. **Section 3.2.2**
    There are examples that show that the evaluation of the change from baseline is almost independent from the actual used measurement method, while the QT measurements themselves show relatively large deviations (Sarapa (2004)). Hence, the recommendation for clinically relevant limits (currently defined by cutoffs of 450, 480 and 500 ms) are of limited value. If kept in the guidance, they should be adaptable to the actual measurement method used, similar to reference ranges which are used as a standard for the evaluation of laboratory safety data.

    Moreover, the absolute cutoffs should be switched with the relative cutoffs (30 and 60 ms), as their relation to a potential risk appears to be closer, because of the independency from the baseline values.

Exceeding the quoted cut-offs in clinical studies should always be reported as "new onset during treatment phase" to take account for potential abnormal baseline conditions.

14. **Section 3.2.3**
    Section 3.2.3 could be removed since there are currently measurement methods available, which provide information on QT dispersion based on QT measurements.

References:

Webber DM, Montague TH, Bird NP, Moss, AJ. Meta-analysis of QTc interval -- pooling data from heterogenous trials. Pharmaceut. Statist., Vol. 1 (2003): 17–23.

Sarapa N, Morganroth J, Couderc JP, Francom SF, Darpo B, Fleishaker JC, McEnroe JD, Chen WT, Zareba W, Moss AJ. Electrographic identification of drug induced QT prolongation: Assessment by different recording and measurement methods. Ann. Noninvas. Electrocard. Vol. 9 (2004): 48-57.